DOCUMENT RESUME

ED 236 197 TM 830 706

AUTHOR Skiba, Russell; And Others
TITLE Characteristics of the Time-Series Data Collected through Curriculum-Based Reading Measurement.
INSTITUTION Minnesota Univ., Minneapolis. Inst. for Research on Learning Disabilities.
SPONS AGENCY Office of Special Education and Rehabilitative Services (ED), Washington, DC.
REPORT NO IRLD-RR-125
PUB DATE Jul 83
CONTRACT 300-80-0622
NOTE 48p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Data Analysis; Diagnostic Teaching; Elementary Education; Reading Tests; *Research Methodology; Research Needs; *Research Problems; Resource Room Programs; Special Education; Visual Discrimination; *Visual Perception
IDENTIFIERS *Data Interpretation; *Time Series Analysis

ABSTRACT

Research on the analysis of time-series data has shown that decisions reached through visual analysis of the data may be influenced by the statistical parameters of those data. The current study investigated the statistical properties of curriculum-based time-series data for 68 elementary resource room students in four Minnesota school districts. Data for slope, standard error of estimate, mean level of performance, and number of data points are presented. Implications for both time-series and conventional pre-post designs are discussed. (Author)

University of Minnesota

Research Report No. 125

CHARACTERISTICS OF THE TIME-SERIES DATA COLLECTED THROUGH CURRICULUM-BASED READING MEASUREMENT

Russell Skiba, Doug Marston, Caren Wesson, Bonita Sevcik
and Stanley L. Deno

IRLD

Institute for Research on Learning Disabilities

Director: James E. Ysseldyke

The Institute for Research on Learning Disabilities is supported by a contract (300-80-0622) with the Office of Special Education, Department of Education, through Title VI-G of Public Law 91-230. Institute investigators are conducting research on the assessment/decision-making/intervention process as it relates to learning disabled students.

During 1980-1983, Institute research focuses on four major areas:

- Referral
- Identification/Classification
- Intervention Planning and Progress Evaluation
- Outcome Evaluation

Additional information on the Institute's research objectives and activities may be obtained by writing to the Editor at the Institute (see Publications list for address).

Research Report No. 125

CHARACTERISTICS OF THE TIME-SERIES DATA COLLECTED THROUGH CURRICULUM-BASED READING MEASUREMENT

Russell Skiba, Doug Marston, Caren Wesson, Bonita Sevcik

and Stanley L. Deno

Institute for Research on Learning Disabilities

University of Minnesota

July, 1983

Abstract

Research on the analysis of time-series data has shown that decisions reached through visual analysis of the data may be influenced by the statistical parameters of those data. The current study investigated the statistical properties of curriculum-based time series data for 68 resource room students in four Minnesota school districts. Data for slope, standard error of estimate, mean level of performance, and number of data points are presented. Implications for both time-series and conventional pre-post designs are discussed.

Characteristics of the Time-Series Data Collected
Through Curriculum-Based Reading Measurement

In their recent volume on single subject research design, Johnston and Pennypacker (1980) pointed out that, in contrast to measurement in the natural sciences, which relies on absolute units of measurement, most research in the social sciences relies on relative units of measurement. As a branch of such research, traditional intelligence and achievement testing is based, not on standard units (i.e., number or duration), but on the relative standing of an individual compared to his/her peers. While such information may be useful for classification and placement decisions (Salvia & Ysseldyke, 1981), it says little about what a student can do on a functionally important task and tends to be relatively insensitive to performance over time (Brown, 1976). Thus, data obtained from traditional assessment provide little or no information regarding individual rates of skill acquisition and are of little assistance in developing or monitoring individual educational programs.

In contrast, the comparatively recent technology of direct and frequent measurement, as exemplified by curriculum-based assessment (Deno & Mirkin, 1977) or precision teaching (White & Haring, 1980), offers a relatively standard unit of measurement, such as words read per minute, for monitoring student progress. Based on single subject research methodology (Sidman, 1960), such systems approach the Johnston and Pennypacker ideal of absolute unit measurement, and provide the teacher with a sensitive recording, over time, of the changes in a child's academic skills. For the researcher, direct and

frequent measurement offers a constant unit, expressed as a rate, that can be used for both between and within subject comparisons. For the teacher, continuous monitoring of skill acquisition allows direct assessment of the impact of various classroom interventions, much as a researcher would test a hypothesis (Deno & Mirkin, 1977).

Kratochwill, Brody, and Piersel (1979) identify a number of advantages that single subject or time-series designs offer for research in the area of learning disabilities. Some time series (such as the ABAB design) allow the researcher to better establish the relationship between the dependent and independent variables. The emphasis on frequent measurement may represent an advantage for both researchers and teachers.

> Learning problems represent a process, dynamic over time, and a research strategy that allows monitoring of these problems, and their relation to an intervention reflects the reality of the practitioner's concern. (Kratochwill et al., 1979, p. 259)

Time-series research designs may offer an alternative strategy in situations where it may not be possible to meet the statistical assumptions necessary to employ conventional pre-post designs. Finally, by focusing on the individual, time-series research designs allow both the researcher and teacher to more clearly specify treatment variables and their relation to subject characteristics, "which leads to refined diagnosis and specific treatment prescription for exceptional children" (Kratochwill et al., 1979, p. 261).

Among those using time-series designs, there has been a tendency to rely on visual analysis of the graphed data. A number of authors point out that such analysis is limited by the technical

characteristics of the data (Jones, Vaught, & Weinrott, 1977: Kazdin, 1976; Parsonson & Baer, 1978). Before an accurate evaluation can be made of the efficacy of intervention, the data must show evidence of stability in the baseline phase. In addition, the variability exhibited in the data will have an influence on conclusions drawn from the data, as will trends exhibited by the data, both before and after interventions. Finally, the number of data points may also determine the nature of conclusions drawn.

The accuracy of judgments made on the basis of visual analysis also may be threatened by serial dependency. Serial dependency refers to the fact that, given repeated measures on the same subject, successive scores are not likely to be independent of one another. The statistical property of serial dependency may have considerable effects on the pattern of data observed (Hartmann, Gottman, Jones, Gardner, Kazdin, & Vaught, 1980). Serial dependency and other statistical properties also may reduce the reliability of visual analysis. Jones, Weinrott, and Vaught (1978) demonstrated that agreement between visual and statistical analyses decreases as serial dependency increases.

Given these difficulties, a number of authors (Gottman & Glass, 1978; Hartmann et al., 1980) have recommended supplementing visual analysis of time-series data with some form of statistical analysis. Procedures have been developed using ANOVA (Gentile, Roden, & Klein, 1972), interrupted time series analysis (Gottman & Glass, 1978), and the c-statistic (Tryon, 1982). In addition, Shine (1975) has proposed a five-step model for integrating single subject and conventional between groups designs.

During the past five years, investigations at the University of Minnesota have established, to a considerable extent, the technical adequacy of using a curriculum-based measurement system based on time series data in the classroom. A series of validity studies (Deno, Mirkin, & Chiang, 1982) showed that reading aloud from a basal reader, reading aloud from lists of isolated words, and guessing the words deleted from a reading passage (i.e., cloze comprehension) all related closely to performance on standardized tests and discriminated between program and grade placement. These formative measures of reading also have shown high test-retest ($\underline{r}$ = .90) and alternate-forms ($\underline{r}$s = .89 -.92) reliability (Shinn, 1981). Finally, both reading from isolated word lists and reading aloud from a basal reader were found to be sensitive to changes within each grade level from fall to spring and across grade levels (Marston, Lowry, Deno, & Mirkin, 1981).

The Minnesota measurement systems include a number of characteristics--absolute unit measurement on a functionally important task, number correct in fixed time, frequent measurement, ability to monitor rate of change over time (slope), and applicability to individual students--that differentiate those systems from traditional assessment. Some data have been collected describing the statistical characteristics of curriculum-based data. Deno, Marston, Mirkin, Lowry, Sindelar, and Jenkins (1982) administered curriculum-based measures to 566 elementary students from three states and described developmental trends in mean level of performance. In a study implementing such measures over a 16-week period, Marston and Deno (1982) reported the mean level of performance and the mean slope of words read correctly across grade and setting.

The purpose of the current study was to investigate further some of the statistical parameters of curriculum-based time-series data. A number of questions can be asked of the data:

1. What is the average increase in reading performance over time? Does the slope representing this increase indicate significant growth in reading performance?

The question of rate of growth is especially crucial for resource room students, for they must show greater growth in order to "catch up" to their peers. While significant weekly improvement in reading may not necessarily preclude demanding still more of students, a failure to show improvement beyond random variability would argue persuasively that the system is not challenging enough.

2. What is the mean level of variability (standard error of estimate) in student performance under a direct and frequent measurement system?

Student performance on any academic task can be expected to vary on a daily basis; the record obtained through curriculum-based measurement obviously will reflect such variability. It is important to know what the average amount of variability is in order to know what constitutes extreme variability.

3. With what frequency can students be expected to meet their goals in a system of direct and frequent measurement?

One would hope that a large percentage of students do in fact meet their goals, indicating both personal success for the student, and instructional success for the teacher.

4. What is the mean level of performance at each grade level?

Although mean level of performance is, in part, determined by the baseline level of performance and long-range goal placement, it still

bears upon the validity of the system. We would in fact expect the measures to discriminate between the performance of children at various grade levels, even if baseline and goal levels were held constant.

Method

Subjects

The subjects were 68 resource room students in three rural and suburban Minnesota school districts. All subjects were participants in research on the effects of teachers using frequent curriculum-based measures. Subjects ranged in grade placement from first to seventh grade; the distribution of students by grade level is shown in Table 1.

Insert Table 1 about here

All students were receiving some resource room instruction and had been receiving such special instruction for anywhere from a few months to six years ($\bar{X}$ = 1.96 years). The time spent in reading instruction in the resource room ranged from 15 minutes to 105 minutes per day, with a mean of 46 minutes per day. The students' teachers averaged two years teaching experience in regular education, and five years in special education.

Procedures

The resource room teachers were trained in the use of the measurement procedures during a series of three half-day workshops at the beginning of the school year. Training was based on the manual,

Procedures to Develop and Monitor Progress on IEP Goals (Mirkin, Deno, Fuchs, Wesson, Tindal, Marston, & Kuehnle, 1981). The teachers continued to use the measures over the entire school year. Visits by observers in December, February, and May, and frequent phone contacts, provided feedback to the teachers on the accuracy of their implementation of the measures.

Measurement consisted of one-minute timed samples of reading from the student's curriculum. Based on the results of previous research, the placement level for testing was set at a criteria of 20-29 words per minute for grades 1 and 2, and 30-39 words per minute for grades 3 through 8. Once this level was determined, passages were chosen randomly from the placement level textbook for measurement purposes. Measurements were conducted three to five times each week. Both number of words read correctly and number of errors in one minute were recorded, and plotted on an equal interval chart. Continuous graphed results allowed teachers to develop a visual record of student progress, like the one represented in Figure 1.

Insert Figure 1 about here

Teachers were instructed to write IEP long-range goals using both the entry level criteria and a desired year-end mastery criteria, usually 70 words correct per minute with no more than 7 errors. The formula used in writing the long-range goal is shown in Figure 2.

Insert Figure 2 about here

Short-term objectives were based on the long-range goals (LRG). In order to compute the short-term objective, teachers first subtracted the baseline level of performance from the criterion level listed in the LRG. Dividing this difference by the number of weeks until the annual review, they arrived at the number of words per week gain necessary to meet the long-range goal criteria. The format used for writing short-term objectives is given in Figure 3.

Insert Figure 3 about here

In addition, the teachers were trained at the beginning of the year, and again at mid-year, in the use of the measurement procedures for evaluation of the instructional program. In order to monitor student growth, the baseline reading level and the long-range goal were connected by an aimline that showed the students' desired progress. Every seven data points, the teachers were to evaluate student growth using a decision rule that required use of the quarter-intersect method (White & Haring, 1980) to determine slope. An example is given in Figure 1. If the student was progressing at a rate equivalent to or greater than that indicated by the aimline, the instructional program was continued; if the projected rate of growth was less than that indicated by the aimline, the teacher was to make a substantial change in the student's program.

Design

Student performance data on direct repeated measures were collected and charted over a six-month period for 68 resource room students. Based on these graphed performance data, a slope for each individual was computed by means of a regression equation. In order to test the probability that the slope represented a significant change in the student's reading performance over time, and not an artifact of individual variability, individual slopes were standardized, and the c-statistic[1] was applied to each individual's data, as recommended by Tryon (1982). Other variables generated from the graphed data were the student's mean reading rate for the entire year, the standard error of estimate (SEE), number of interventions, and total number of data points. Finally, two measures were used to determine whether students achieved their reading goals: first, whether any data points fell above the long-range goal, and second, the number of data points that fell above the level specified for the long-range goal.

Results

The average number of measurements per week ranged from near 0 to 4.7, with a mean of 2.8 data points per week. The total number of data points over the six-month period ranged from 20 to 131, with a mean of 51.8. The long-range goal for those in grades 3-7 averaged 72 words per minute, while for those in grades 1 and 2 the mean LRG was 53 words per minute.

The mean slope (average number of words gained per week) for all students was 1.55 words per minute gain per week. The average number

of words gained per week tended to be greater the lower the grade level (see Table 2). As noted previously, the c-statistic was applied to all slopes: conversion of the c-statistic into a z-score[2] (Tryon, 1982) allows one to determine the .05 level of significance. The large majority of students exhibited slopes that represented significant gains in words read correctly. Only 22% of the students' slopes were insignificant (see Table 3).

Insert Tables 2 and 3 about here

The mean standard error of estimate, a measure of variability, was 10.17 words. This means that student data varied, on the average, ±5.09 words per minute around the slope. The SEE's ranged from 8.45 words (grade 2) to 11.56 words (grade 4) (see Table 4). The amount of variability in the graphed data showed an increase from second to third grade, then tended to level off for grades 3 through 5. Table 5 presents average and high levels of variability for first and second, and for third through sixth grades.

Insert Tables 4 and 5 about here

The distribution of mean number of words read correctly by grade is presented in Table 6. Although mean number of words read correctly for any given student is determined in large part by the levels at which the baseline and long-range goal are set for that student, the mean performance did tend to increase with grade level.

Insert Table 6 about here

As can be seen in Table 7, students had little difficulty meeting their long-range goals. Over 85% did so at least once. In fact, the mean number of data points falling at or above the long-range goal level was 8.2. Thus, many students equaled or exceeded their goal a number of times over the course of the six-month period.

Insert Table 7 about here

Although teachers were trained in evaluating the data, and were encouraged to make changes in the students' educational program if the data warranted, few interventions were implemented by teachers (see Table 8). The mean number of program changes implemented per student was .65 for the entire school year.

Insert Table 8 about here

Student Time Series Data

Examples of individual student time series are presented in Figures 4-7; these were chosen from actual student data to provide clear illustrations of statistical properties. Regression slope, standard error of estimate, mean for the year, and total number of data points are reported for each graph and may be compared with the sample means in Tables 2, 4, 5 and 6. Connected points represent

words read correctly from text in one minute, X's represent words read incorrectly in the same time interval. The first three to six data points before the vertical line represent baseline sampling to determine level of placement for measurement; due to the initial sampling procedure described above, some of the baseline measurements may be higher or lower than post-baseline measurement. The diagonal drawn from the baseline to the long-range goal represents the aimline set by the teacher at the beginning of the year. This aimline may or may not reflect the actual rate of student progress as expressed in the slope.

Insert Figures 4-7 about here

Note that individual variations in rate of acquisition, variability, and mean level of performance preclude description of student time series data by any single statistical property. Both Figure 4 and Figure 5 show relatively steep slopes, yet differ greatly in the variability around those slopes. The standard errors of estimate for Figure 5 and 6 are more similar, but Figure 6 shows a much flatter slope. Finally, although the mean levels of responding are very similar for the data presented in Figures 4 and 7, the rate of acquisition differs greatly.

Discussion

The curriculum-based measurement system described herein provides teachers with a method of making curriculum decisions based on analysis of time series data (Deno & Mirkin, 1977). Since both

training in statistical analysis and access to statistical programs may be limited for teachers, it is likely that decisions about the data will be based primarily on visual analysis. Given the concerns expressed regarding the limitations of visual analysis of time series data, it is important that the statistical properties that determine the accuracy of visual inference be thoroughly explored. The present study has investigated statistical characteristics such as trend, variability, level of performance, and number of data points in time-series data obtained from resource room students.

Perhaps the most important property of time series data that directly influences visual analysis is the trend of the data. The slopes presented for grades 2-5 are very similar to those obtained by Marston and Deno (1982) for resource room students and, like that sample, tend to be flatter than the slopes obtained for students in either regular education or Title I programs in the Marston and Deno sample. Caution must be used in comparing across different samples. Nevertheless, given that the purpose of special education is to accelerate educational progress so that students in such settings may return to the mainstream, these results could be viewed as rather discouraging.

Few interventions, however, were made in students' educational programs to accelerate the slopes. The slopes obtained from this sample may thus represent a baseline level of performance that could be accelerated given greater attention to planned interventions. It is interesting to note that most behavioral literature views the ideal stable baseline as one with minimum acceleration or deceleration

(Hersen & Barlow, 1976). In tasks such as reading, however, in which we expect continuous acquisition (at least until high school), lack of acceleration may be neither necessary nor desirable. Many students may, in fact, be referred for special placement because they fail to exhibit the continuous acquisition characteristic of most elementary students. The slopes obtained for this sample may thus represent a baseline level of rate of acquisition that is likely to be lower than the rate of acquisition for students in regular education. In such a case, the purpose of interventions is not to increase or decrease the mean level of responding in relation to a stable baseline, but rather to increase the rate of acquisition (slope) in relation to the current rate.

Decisions about when to implement such interventions may be highly influenced by the degree of variability of the data (Hersen & Barlow, 1976). The standard error of estimate data presented here thus may be critically important in interpreting the data. The current results, if validated by continued study, may provide "norms" of moderate and high variability that could be used to guide teachers in decision making. A high degree of variability in student scores might provide a caution in making decisions based on those data, and might indicate the need to bring the variability under control before planning educational interventions. On the other hand, a well planned intervention might in itself reduce the rate of variability.

Teachers in the current sample seemed aware that external or historical events could affect the variability of student performance. A number of teachers noted that increased variability often seemed to

coincide with reports of increased instability in the student's home situation. The teacher's comments concerning the time series data presented in Figure 7 represent that teacher's hypotheses concerning the causes of variability. The fact that time series data encourage formulation of hypotheses that may then be tested represents another advantage of the time-series design over conventional pre-post measurement.

As mean level of performance was controlled, to a certain extent, by level of curriculum placement (i.e., all students initially were placed in material in which they read 20-39 words correctly per minute) one would expect to detect few developmental trends in the data. Nonetheless, though not statistically significant, certain trends do appear. Mean level of performance tends to increase by grade level, while rate of acquisition as represented by the slope tends to decrease from grade 2 through grade 5. These results tend to support the findings of Deno et al. (1982) that growth on academic time series measures can be represented by a curve that is negatively accelerated; that is, rate of acquisition is greatest across grades 1, 2, and 3 and tends to accelerate at a slower rate for students in grades 4, 5 and 6.

The ideal rate of academic growth has yet to be determined in special education. Lindsley (1982) argues that it is better to set higher goals and demand steeper learning slopes and that, by setting goals too low, we are in danger of training the child to meet our lower expectations. On the other hand, goals set unrealistically high may be unattainable for the student in a resource room setting and may only add to that student's experience of frustration in education.

The majority of students in the present analysis reached their goals at least once, and the number of data points at or above the long-range goal level ($\bar{X} = 8.2$) indicates that the goal was, for the most part, solidly attained. The goals set by curriculum-based measurement thus seem to be eminently reachable to the majority of resource room students. In addition, the significance of the c-scores for the majority of slopes indicates that the obtained slopes represent real trends for these students, and not merely random variation.

The current study has begun to delimit some of the statistical characteristics of curriculum-based time series data. Given the relatively low rate of acquisition represented by the slopes for this sample, one would hope that such characteristics are modifiable. Further research is needed to determine whether it is possible to accelerate rate of acquisition either by manipulation of goals and aimlines, or by increasing the rate of teachers' interventions.

The findings of this study join previous literature (Gottman & Glass, 1978; Hartmann et al., 1980) in arguing for both visual and statistical analysis of time-series data. It is especially apparent from an examination of individual data that no single statistical property of time-series data can be used to rank order individuals, since individual students may vary on a number of such dimensions simultaneously. Such a finding has implications beyond time-series analysis, since standardized testing generally rank orders students on the basis of a mean score at one point in time. Analysis of time-series data suggests that two students with the same mean may have

arrived at that mean through very different rates of acquisition. Thus, curriculum-based time-series measurement may provide a more accurate and complete record of student educational progress.

References

Brown, F. G. Principles of educational and psychological testing (2nd ed.). New York: Holt, Rinehart, & Winston, 1976.

Deno, S., Marston, D., Mirkin, P., Lowry, L., Sindelar, P., & Jenkins, J. The use of standard tasks to measure achievement in reading, spelling, and written expression: A normative and developmental study (Research Report No. 87). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1982.

Deno, S., & Mirkin, P. Data-based program modification: A manual. Minneapolis: Leadership Training Institute, 1977.

Deno, S. L., Mirkin, P. K., & Chiang, B. Identifying valid measures of reading. Exceptional Children, 1982, 49, 36-47.

Gentile, J. R., Roden, A. H., & Klein, R. D. An analysis of variance model for the intrasubject replication design. Journal of Applied Behavior Analysis, 1972, 5, 193-198.

Gottman, J. M., & Glass, G. V. Analysis of interrupted time-series experiments. In T. R. Kratochwill (Ed.), Single-subject research: Strategies for evaluating change. New York: Academic Press, 1978.

Hartmann, D. P., Gottman, J. M., Jones, R. R., Gardner, W., Kazdin, A. E., & Vaught, R. Interrupted time-series analysis and its application to behavioral data. Journal of Applied Behavior Analysis, 1980, 13, 543-559.

Hersen, M., & Barlow, D. H. Single case experimental designs: Strategies for studying behavior change. New York: Pergamon Press, 1976.

Johnston, J. M., & Pennypacker, H. S. Strategies and tactics of human behavioral research. Hillsdale, NJ: Lawrence Erlbaum Associates, 1980.

Jones, R. R., Vaught, R. S., & Weinrott, M. Time-series analysis in operant research. Journal of Applied Behavior Analysis, 1977, 10, 151-166.

Jones, R. R., Weinrott, M. R., & Vaught, R. S. Effects of serial dependency on the agreement between visual and statistical inference. Journal of Applied Behavior Analysis, 1978, 11, 277-283.

Kazdin, A. E. Statistical analysis for single-case experimental designs. In M. Hersen & D. Barlow (Eds.), Single case experimental designs: Strategies for studying behavior change, New York: Pergamon Press, 1976.

Kratochwill, T. R., Brody, G. H., & Piersel, W. C. Time-series research: Some comments on design methodology for research in learning disabilities. Journal of Learning Disabilities, 1979, 12(4), 257-263.

Lindsley, O. R. How teachers could be rewarded for student learning. Paper presented at the Twelfth Annual Invitational Conference on Leadership in Special Education, Minneapolis, November, 1982.

Marston, D., & Deno, S. L. Implementation of direct and repeated measurement in the school setting (Research Report No. 106). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1982.

Marston, D., Lowry, L., Deno, S., & Mirkin, P. An analysis of learning trends in simple measures of reading, spelling and written expression: A longitudinal study (Research Report No. 49). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1981.

Mirkin, P., Deno, S., Fuchs, L., Wesson, C., Tindal, G., Marston, D., & Kuehnle, K. Procedures to develop and monitor progress on IEP goals. Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1981.

Parsonson, B. S., & Baer, D. M. The analysis and presentation of graphic data. In T. R. Kratochwill (Ed.) Single subject research: Strategies for evaluating change. New York: Academic Press, 1978.

Salvia, J., & Ysseldyke, J. Assessment in special and remedial education. Boston: Houghton Mifflin, 1981.

Shine, L. C. Five research steps designed to integrate the single-subject and multi-subject approaches to experimental research. Canadian Psychological Review, 1975, 16, 179-184.

Shinn, M. A comparison of psychometric and functional differences between students labeled learning disabled and low achieving. Unpublished doctoral dissertation, University of Minnesota, 1981.

Sidman, M. Tactics of scientific research. New York: Basic Books, 1960.

Tryon, W. W. A simplified time-series analysis for evaluating treatment interventions. Journal of Applied Behavior Analysis, 1982, 15, 423-429.

White, O. R., & Haring, N. G. Exceptional teaching (2nd ed.). Columbus, OH: Charles E. Merrill, 1980.

Footnotes

The authors gratefully acknowledge the contributions of Dr. Gerald Tindal in directing our attention to the c-statistic as a method of time-series analysis.

[1] The actual calculations involved in the c-statistic, as cited by Tryon (1982), are:

$$C = 1 - \frac{\sum_{i=1}^{N-1} (X_i - X_{i-1})^2}{Z\sum_{i=1}^{N} (X_i - \overline{X})^2}$$

where the numerator of the right hand term is the sum of the (N-1) squared consecutive differences associated with the time series. The denominator is twice the sum of the (N) squared deviations of the time-series data points from their norm.

[2] The standard error of the c-statistic is

$$Sc = \sqrt{\frac{N+2}{(N-1)(N+1)}}$$

The c-statistic may be converted to a z-statistic and tested for significance through the following ratio:

$$Z = \frac{C}{Sc}$$

Table 1

Distribution of Students by Grade Level

Grade	Number of Students	Percentage
1	2	2.9
2	18	26.5
3	15	22.1
4	16	23.5
5	12	17.6
6	3	4.4
7	2	2

Table 2

Mean Slope or Average Number of Words Gained Per Week[a]

Grade	Mean	Standard Deviation
2	1.78	1.21
3	1.62	.71
4	1.42	.92
5	1.36	.66

[a]Means and standard deviations are not reported for grades 1, 6, 7, and 8 due to low N's.

Table 3

Students with Significant Increasing Slopes

	Frequency	Percentage
Students with insignificant slopes	15	22%
Students with significant slopes	53	78%

Table 4

Mean Standard Error of Estimate at Each Grade Level

Grade	Mean	Standard Deviation
2	8.45	2.27
3	10.06	4.29
4	11.56	5.39
5	10.35	4.43

Table 5

Standard Error of Estimate

Normal and High Variability

	Mean	Median	Percentile 75th	90th
Grades 1 and 2	8.36	8.24	9.91	10.63
Grades 3 through 6	10.65	9.36	13.97	16.51

Table 6

Mean Number of Words Read Correctly

Grade	Mean	Standard Deviation
2	42.24	10.59
3	51.23	8.71
4	54.13	8.97
5	54.27	10.88

Table 7

Number of Students Reaching Long-Range Goal at Least Once

	Frequency	Percentage
Students not reaching long-range goal	10	14.7%
Students reaching long-range goal at least once	58	85.3%

Table 8

Number of Interventions Implemented in

Student Instruction Programs

Number of Changes	Frequency
0	45
1	12
2	4
3	4
4	3

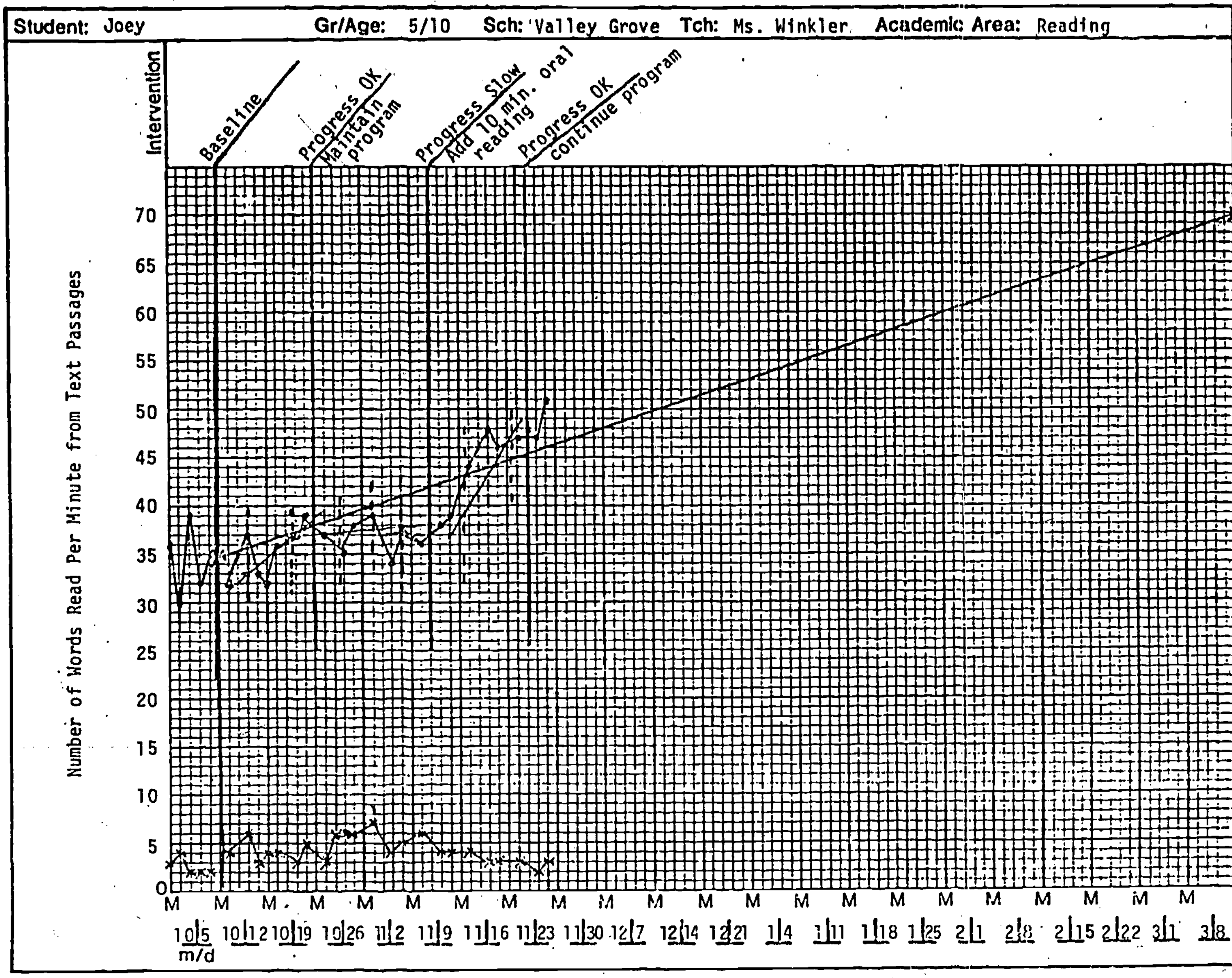

Figure 1. Example of Individual Time-Series Data.

	Condition	Behavior	Criteria
LRG:	In ________ (total # weeks) weeks, when presented with stories from Level ____ (#), ________ (reading series),	student will read aloud	at the rate of 50 wpm or better 5 or fewer errors.

Figure 2. Format for Long-Range Goal: Reading

	Condition	Behavior	Criteria
STO	Each successive week, when presented with a random selection from Level ____ (#) of ____ (reading series),	student will read aloud	at an average increase of ____ (repeated-actual performance/total # weeks) words correct/minute and no increase in errors.

Figure 3. Format for Short-Term Objective: Reading

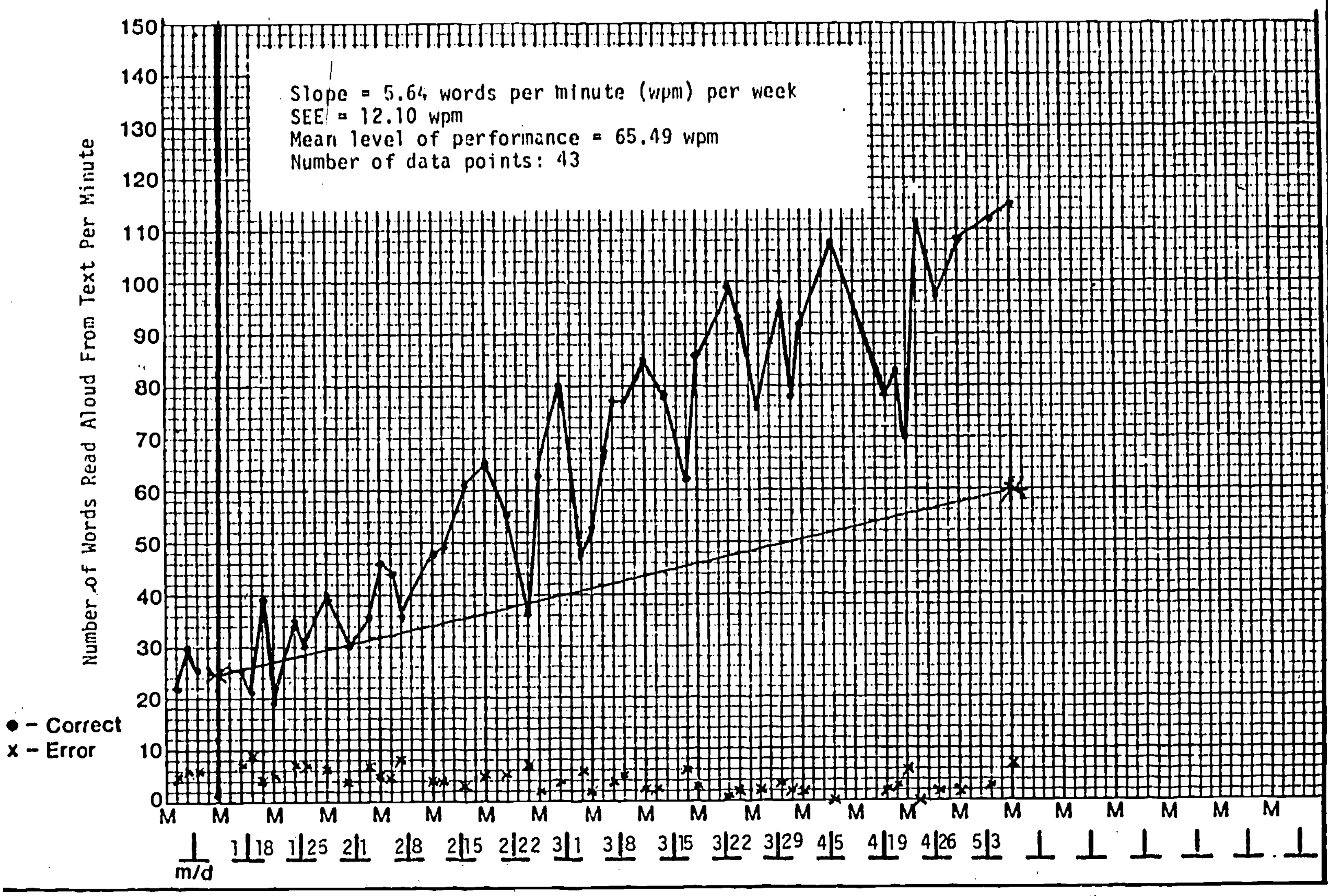

Figure 4. Example of time-series for student in grade 2.

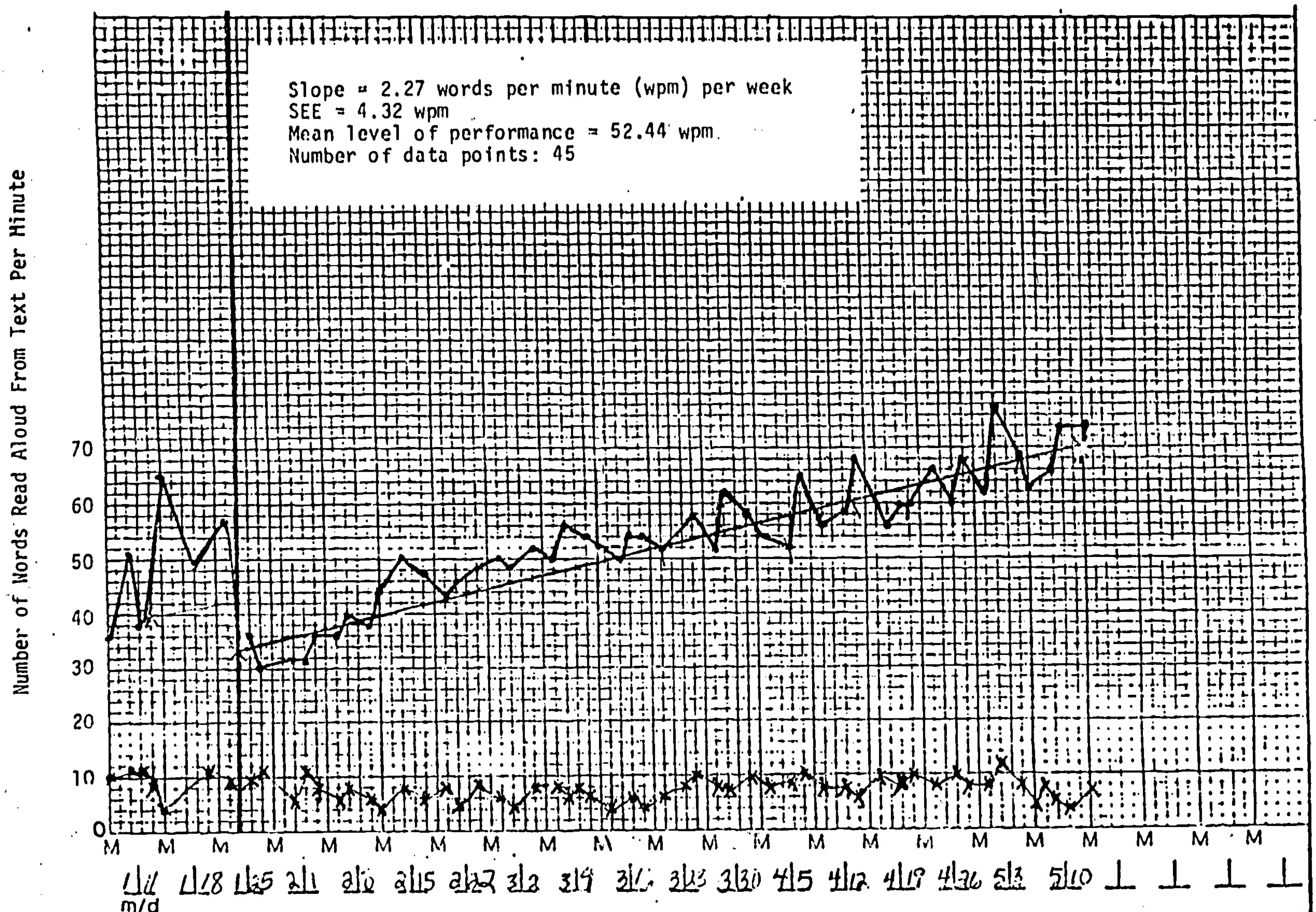

Figure 5. Example of time-series for student in grade 4.

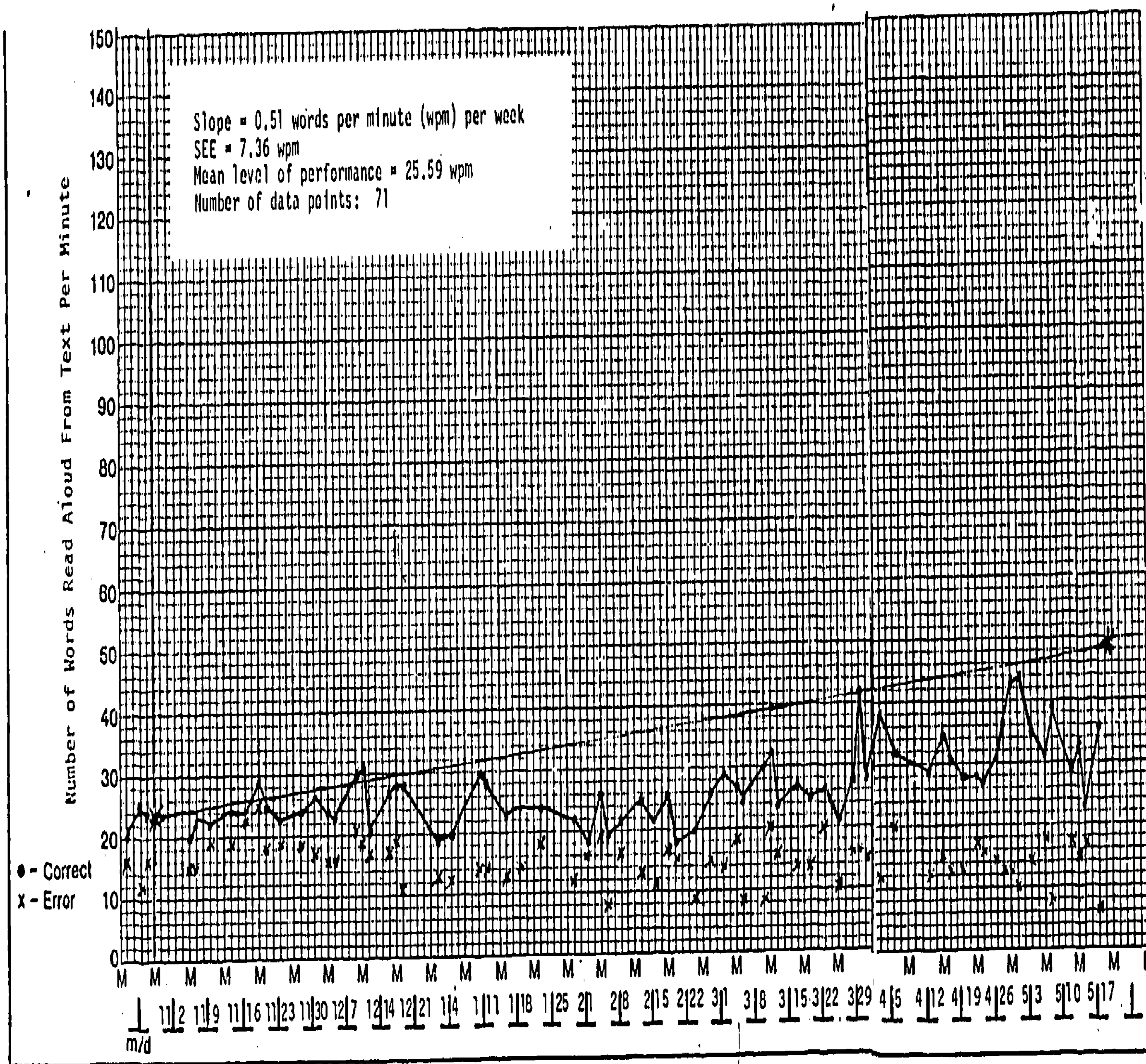

Figure 6. Example of time-series for student in grade 2.

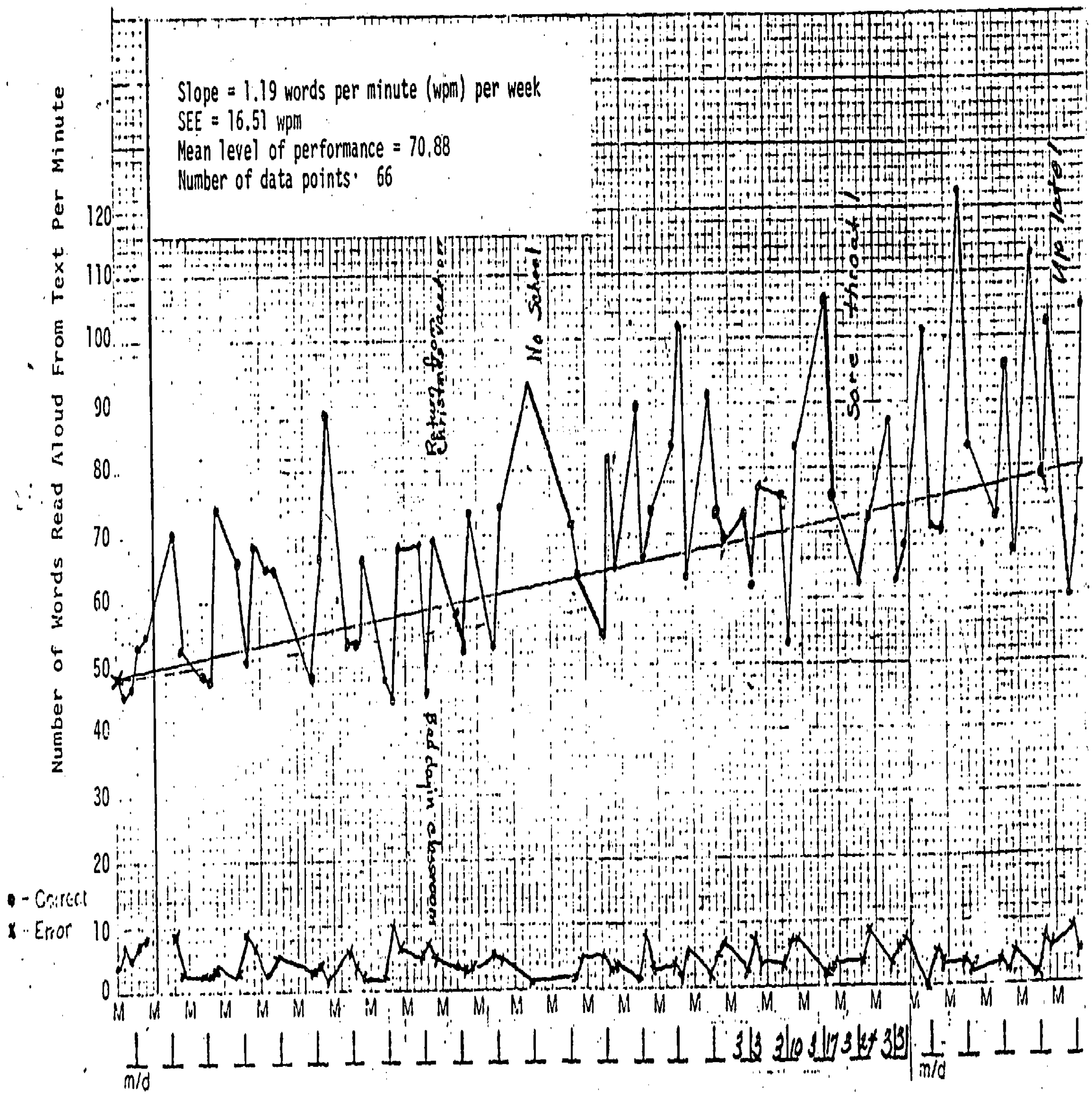

Figure 7. Example of time-series for student in grade 5.

35

PUBLICATIONS

Institute for Research on Learning Disabilities
University of Minnesota

The Institute is not funded for the distribution of its publications. Publications may be obtained for $4.00 each, a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be pre-paid. Requests should be directed to: Editor, IRLD, 350 Elliott Hall; 75 East River Road, University of Minnesota, Minneapolis, MN 55455.

The publications listed here are only those that have been prepared since 1982. For a complete, annotated list of all IRLD publications, write to the Editor.

Wesson, C., Mirkin, P., & Deno, S. Teachers' use of self instructional materials for learning procedures for developing and monitoring progress on IEP goals (Research Report No. 63). January, 1982.

Fuchs, L., Wesson, C., Tindal, G., Mirkin, P., & Deno, S. Instructional changes, student performance, and teacher preferences: The effects of specific measurement and evaluation procedures (Research Report No. 64). January, 1982.

Potter, M., & Mirkin, P. Instructional planning and implementation practices of elementary and secondary resource room teachers: Is there a difference? (Research Report No. 65). January, 1982.

Thurlow, M. L., & Ysseldyke, J. E. Teachers' beliefs about LD students (Research Report No. 66). January, 1982.

Graden, J., Thurlow, M. L., & Ysseldyke, J. E. Academic engaged time and its relationship to learning: A review of the literature (Monograph No. 17). January, 1982.

King, R., Wesson, C., & Deno, S. Direct and frequent measurement of student performance: Does it take too much time? (Research Report No. 67). February, 1982.

Greener, J. W., & Thurlow, M. L. Teacher opinions about professional education training programs (Research Report No. 68). March, 1982.

Algozzine, B., & Ysseldyke, J. Learning disabilities as a subset of school failure: The oversophistication of a concept (Research Report No. 69). March, 1982.

Fuchs, D., Zern, D. S., & Fuchs, L. S. A microanalysis of participant behavior in familiar and unfamiliar test conditions (Research Report No. 70). March, 1982.

Shinn, M. R., Ysseldyke, J., Deno, S., & Tindal, G. A comparison of psychometric and functional differences between students labeled learning disabled and low achieving (Research Report No. 71). March, 1982.

Thurlow, M. L. Graden, J., Greener, J. W., & Ysseldyke, J. E. Academic responding time for LD and non-LD students (Research Report No. 72). April, 1982.

Graden, J., Thurlow, M., & Ysseldyke, J. Instructional ecology and academic responding time for students at three levels of teacher-perceived behavioral competence (Research Report No. 73). April, 1982.

Algozzine, B., Ysseldyke, J., & Christenson, S. The influence of teachers' tolerances for specific kinds of behaviors on their ratings of a third grade student (Research Report No. 74). April, 1982.

Wesson, C., Deno, S., & Mirkin, P. Research on developing and monitoring progress on IEP goals: Current findings and implications for practice (Monograph No. 18). April, 1982.

Mirkin, P., Marston, D., & Deno, S. L. Direct and repeated measurement of academic skills: An alternative to traditional screening, referral, and identification of learning disabled students (Research Report No. 75). May, 1982.

Algozzine, B., Ysseldyke, J., Christenson, S., & Thurlow, M. Teachers' intervention choices for children exhibiting different behaviors in school (Research Report No. 76). June, 1982.

Tucker, J., Stevens, L. J., & Ysseldyke, J. E. Learning disabilities: The experts speak out (Research Report No. 77). June, 1982.

Thurlow, M. L., Ysseldyke, J. E., Graden, J., Greener, J. W., & Mecklenberg, C. Academic responding time for LD students receiving different levels of special education services (Research Report No. 78). June, 1982.

Graden, J. L., Thurlow, M. L., Ysseldyke, J. E., & Algozzine, B. Instructional ecology and academic responding time for students in different reading groups (Research Report No. 79). July, 1982.

Mirkin, P. K., & Potter, M. L. A survey of program planning and implementation practices of LD teachers (Research Report No. 80). July, 1982.

Fuchs, L. S., Fuchs, D., & Warren, L. M. Special education practice in evaluating student progress toward goals (Research Report No. 81). July, 1982.

Kuehnle, K., Deno, S. L., & Mirkin, P. K. Behavioral measurement of social adjustment: What behaviors? What setting? (Research Report No. 82). July, 1982.

Fuchs, D., Dailey, Ann Madsen, & Fuchs, L. S. Examiner familiarity and the relation between qualitative and quantitative indices of expressive language (Research Report No. 83). July, 1982.

Videen, J., Deno, S., & Marston, D. Correct word sequences: A valid indicator of proficiency in written expression (Research Report No. 84). July, 1982.

Potter, M. L. Application of a decision theory model to eligibility and classification decisions in special education (Research Report No. 85). July, 1982.

Greener, J. E., Thurlow, M. L., Graden, J. L., & Ysseldyke, J. E. The educational environment and students' responding times as a function of students' teacher-perceived academic competence (Research Report No. 86). August, 1982.

Deno, S., Marston, D., Mirkin, P., Lowry, L., Sindelar, P., & Jenkins, J. The use of standard tasks to measure achievement in reading, spelling, and written expression: A normative and developmental study (Research Report No. 87). August, 1982.

Skiba, R., Wesson, C., & Deno, S. L. The effects of training teachers in the use of formative evaluation in reading: An experimental-control comparison (Research Report No. 88). September, 1982.

Marston, D., Tindal, G., & Deno, S. L. Eligibility for learning disability services: A direct and repeated measurement approach (Research Report No. 89). September, 1982.

Thurlow, M. L., Ysseldyke, J. E., & Graden, J. L. LD students' active academic responding in regular and resource classrooms (Research Report No. 90). September, 1982.

Ysseldyke, J. E., Christenson, S., Pianta, R., Thurlow, M. L., & Algozzine, B. An analysis of current practice in referring students for psychoeducational evaluation: Implications for change (Research Report No. 91). October, 1982.

Ysseldyke, J. E., Algozzine, B., & Epps, S. A logical and empirical analysis of current practices in classifying students as handicapped (Research Report No. 92). October, 1982.

Tindal, G., Marston, D., Deno, S. L., & Germann, G. Curriculum differences in direct repeated measures of reading (Research Report No. 93). October, 1982.

Fuchs, L.S., Deno, S. L., & Marston, D. Use of aggregation to improve the reliability of simple direct measures of academic performance (Research Report No. 94). October, 1982.

Ysseldyke, J. E., Thurlow, M. L., Mecklenburg, C., & Graden, J. Observed changes in instruction and student responding as a function of referral and special education placement (Research Report No. 95). October, 1982.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. Effects of frequent curriculum-based measurement and evaluation on student achievement and knowledge of performance: An experimental study (Research Report No. 96). November, 1982.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. Direct and frequent measurement and evaluation: Effects on instruction and estimates of student progress (Research Report No. 97). November, 1982.

Tindal, G., Wesson, C., Germann, G., Deno, S. L., & Mirkin, P. K. The Pine County model for special education delivery: A data-based system (Monograph No. 19). November, 1982.

Epps, S., Ysseldyke, J. E., & Algozzine, B. An analysis of the conceptual framework underlying definitions of learning disabilities (Research Report No. 98). November, 1982.

Epps, S., Ysseldyke, J. E., & Algozzine, B. Public-policy implications of different definitions of learning disabilities (Research Report No. 99). November, 1982.

Ysseldyke, J. E., Thurlow, M. L., Graden, J. L., Wesson, C., Deno, S. L., & Algozzine, B. Generalizations from five years of research on assessment and decision making (Research Report No. 100). November, 1982.

Marston, D., & Deno, S. L. Measuring academic progress of students with learning difficulties: A comparison of the semi-logarithmic chart and equal interval graph paper (Research Report No. 101). November, 1982.

Beattie, S., Grise, P., & Algozzine, B. Effects of test modifications on minimum competency test performance of third grade learning disabled students (Research Report No. 102). December, 1982

Algozzine, B., Ysseldyke, J. E., & Christenson, S. An analysis of the incidence of special class placement: The masses are burgeoning (Research Report No. 103). December, 1982.

Marston, D., Tindal, G., & Deno, S. L. Predictive efficiency of direct, repeated measurement: An analysis of cost and accuracy in classification (Research Report No. 104). December, 1982.

Wesson, C., Deno, S., Mirkin, P., Sevcik, B., Skiba, R., King, R., Tindal, G., & Maruyama, G. Teaching structure and student achievement effects of curriculum-based measurement: A causal (structural) analysis (Research Report No. 105). December, 1982.

Mirkin, P. K., Fuchs, L. S., & Deno, S. L. (Eds.). Considerations for designing a continuous evaluation system: An integrative review (Monograph No. 20). December, 1982.

Marston, D., & Deno, S. L. Implementation of direct and repeated measurement in the school setting (Research Report No. 106). December, 1982.

Deno, S. L., King, R., Skiba, R., Sevcik, B., & Wesson, C. The structure of instruction rating scale (SIRS): Development and technical characteristics (Research Report No. 107). January, 1983.

Thurlow, M. L., Ysseldyke, J. E., & Casey, A. Criteria for identifying LD students: Definitional problems exemplified (Research Report No. 108). January, 1983.

Tindal, G., Marston, D., & Deno, S. L. The reliability of direct and repeated measurement (Research Report No. 109). February, 1983.

Fuchs, D., Fuchs, L. S., Dailey, A. M., & Power, M. H. Effects of pre-test contact with experienced and inexperienced examiners on handicapped children's performance (Research Report No. 110). February, 1983

King, R. P., Deno, S., Mirkin, P., & Wesson, C. The effects of training teachers in the use of formative evaluation in reading: An experimental-control comparison (Research Report No. 111). February, 1983.

Tindal, G., Deno, S. L., & Ysseldyke, J. E. Visual analysis of time series data: Factors of influence and level of reliability (Research Report No. 112). March, 1983.

Tindal, G, Shinn, M., Fuchs, L., Fuchs, D., Deno, S., & Germann, G. The technical adequacy of a basal reading series mastery test (Research Report No. 113). April, 1983.

Sevcik, B., Skiba, R., Tindal, G., King, R., Wesson, C., Mirkin, P., & Deno, S. Communication of IEP goals and student progress among parents, regular classroom teachers, and administrators using systematic formative evaluation (Research Report No. 114). April, 1983.

Wesson, C. Two student self-management techniques applied to data-based program modification (Research Report No. 115). April, 1983.

Wesson, C., Skiba, R., Sevcik, B., King, R., Tindal, G., Mirkin, P., & Deno, S. The impact of the structure of instruction and the use of technically adequate instructional data on reading improvement (Research Report No. 116). May, 1983.

Wesson, C. Teacher vs student selection of instructional activities (Research Report No. 117). May, 1983.

Tindal, G., & Deno, S. Factors influencing the agreement between visual and statistical analyses of time series data (Research Report No. 118). June, 1983.

Skiba, R. S. Classroom behavior management: A review of the literature (Monograph No. 21), June, 1983.

Graden, J. L., Thurlow, M. L., & Ysseldyke, J. E. When are students most academically engaged? Academic responding time in different instructional ecologies (Research Report No. 119). June, 1983.

Fuchs, L. S., Deno, S. L., & Roettger, A. The effect of alternative data-utilization rules on spelling achievement: An n of 1 study (Research Report No. 120). June, 1983.

Skiba, R., Sevcik, B., Wesson, C., King, R., & Deno, S. The non-effect of process-product variables in resource classrooms (Research Report No. 121). June, 1983.

Fuchs, L. Tindal, G., Fuchs, D., Deno, S., & Germann, G. Technical adequacy of basal readers' mastery tests: The Ginn 720 series (Research Report No. 122). June, 1983.

Tindal, G., Germann, G., Marston, D., & Deno, S. The effectiveness of special education: A direct measurement approach (Research Report No. 123). June, 1983.

Sevcik, B., Skiba, R., Tindal, G., King, R., Wesson, C., Mirkin, P., & Deno, S. Curriculum-based measurement: Effects on instruction, teacher estimates of student progress, and student knowledge of performance (Research Report No. 124). July, 1983.

Skiba, R., Marston, D., Wesson, C., Sevcik, B., & Deno, S. L. Characteristics of the time-series data collected through curriculum-based reading measurement (Research Report No. 125). July, 1983.

Ysseldyke, J., Christenson, S., Graden, J., & Hill, D. Practical implications of research on referral and opportunity to learn (Monograph No. 22). July, 1983.

Marston, D., Deno, S., & Tindal, G. A comparison of standardized achievement tests and direct measurement techniques in measuring pupil progress (Research Report No. 126). July, 1983.

Fuchs, D., Fuchs, L. S., Tindal, G., & Deno, S. L. Variability of performance: A "signature" characteristic of learning disabled children? (Research Report No. 127). July, 1983.